JHR 94: 215–246 (2022)
doi: 10.3897/jhr.94.84617
https://jhr.pensoft.net

RESEARCH ARTICLE

JOURNAL OF Hymenoptera RESEARCH
A peer-reviewed open-access journal
The International Society of Hymenopterists

A metabarcoding framework for wild bee assessment in Luxembourg

Fernanda Herrera-Mesías[1,2], Imen Kharrat Ep Jarboui[1], Alexander M. Weigand[1]

1 *Musée National d'Histoire Naturelle de Luxembourg, Zoology Department, Luxembourg* **2** *Ruhr-Universität, Department of Animal Ecology, Evolution and Biodiversity, Bochum, Germany*

Corresponding author: Alexander M. Weigand (alexander.weigand@mnhn.lu)

Academic editor: Michael Ohl | Received 30 March 2022 | Accepted 19 October 2022 | Published 20 December 2022

https://zoobank.org/1F2CA60D-E91C-4B05-98A5-A6C29C78B114

Citation: Herrera-Mesías F, Ep Jarboui IK, Weigand AM (2022) A metabarcoding framework for wild bee assessment in Luxembourg. Journal of Hymenoptera Research 94: 215–246. https://doi.org/10.3897/jhr.94.84617

Abstract

Wild bees are crucial organisms for terrestrial environments. Their ongoing decline could cause irreparable damage to ecosystem services vital to plant reproduction and human food production. The importance of taking swift action to prevent further declines is widely acknowledged, but the current deficit of reliable taxonomic information complicates the development of efficient conservation strategies targeting wild bees. DNA metabarcoding can help to improve this situation by providing rapid and standardized mass identification. This technique allows the analysis of large numbers of specimens without the need for specialized taxonomic knowledge by matching high-throughput sequencing reads against public DNA barcode reference libraries. However, the validation of this approach for wild bees requires the evaluation of potential error sources on a regional scale. Here we analyzed the effects of three potential error sources on a metabarcoding pipeline customized for the wild bee fauna of Luxembourg. In an *in silico* study, we checked the completeness of the BOLD reference library for 349 species found in the country, the correspondence between molecular and morphological species delimitation for these taxa, and the amplification efficiency of three commonly used metabarcoding primer pairs (mlCOIintF/HCO2198, LepF1/MLepF1-Rev and BF2/BR2). The detection power of the pipeline was evaluated based on the species recovery rates from mock communities of known composition under variable DNA concentration treatments. The reference barcode library evaluation results show that 97% of the species have at least a single barcode in BOLD Systems (minimal length 196 bp) and that 85% of species have ≥ 5 barcodes in the public domain. The mlCOIintF/HCO2198 target fragment presented the highest coverage (77.94% of the species with full barcode sequences), followed by the target fragments of LepF1/MLepF1-Rev (77.65%) and BF2/BR2 (68.48%). Only 60% of the morphospecies presented a complete coverage of the prominent Folmer region (658 bp). The *in silico* amplification efficiency analysis shows that the BF2/BR2 primer pair has the best-predicted amplification performance, but none of the primer combinations evaluated can be expected to efficiently amplify all local wild bee genera. Finally, all species detection rates in the mock communities, except for the sample with the most discrepant

DNA concentrations, were above 97%, with no significant differences found among treatments. These results indicate that the detection capacity of the pipeline is robust enough to be used for the reliable assessment of local wild bee biodiversity, even if species from various size categories are pooled together. Primer bias has a major effect on species detection, which can be acknowledged with a preliminary assessment of primer-template mismatch and sophisticated methodological designs (e.g. mock community controls, replicates). Overall, the metabarcoding pipeline here described provides a suitable tool for quick and reliable taxonomic identification of the regional wild bee fauna to aid conservation initiatives in Luxembourg – and beyond.

Keywords
biodiversity, COI, conservation, molecular taxonomic tools, pollinators, primer evaluation

Introduction

Bees (Hymenoptera, Anthophila) are important insect pollinators of Angiosperms with critical ecological functions and high economic value (Brown and Paxton 2009). Their pollination services have a crucial impact on the reproduction of both wild flowering plants and cultivated crops (Potts et al. 2010a; Rafferty 2017). Over 75% of the crops grown worldwide benefit to some degree from insect-mediated pollination and wild bees participate directly in the reproduction of about 42% of the leading food crops grown for human consumption (Klein et al. 2007; Potts et al. 2010a). Moreover, the global annual economic value of insect pollination, most of which is performed by bees, has been estimated on at least 153 billion euros (Potts et al. 2010a).

Despite the high importance of the ecological services provided by insects (Losey and Vaughan 2006), several studies have reported large declines in insect diversity, abundance and biomass over the past few decades (Dirzo et al. 2014; Hallmann et al. 2017; Hausmann et al. 2020), a trend from which bees are not excluded. Consistent evidence has been found of ongoing decline in Europe for honey bees and bumblebees (Rasmont and Mersch 1988; Goulson et al. 2008; Potts et al. 2010a, b). Rarefaction analyses performed on records from national entomological databases in the UK and The Netherlands suggest that wild bee biodiversity has significantly declined after 1980 in landscapes of both countries, with a special emphasis on species with narrow habitat requirements (Biesmeijer et al. 2006). Similar trends are likely to be true for nearby regions and similarly urbanized locations, but important documentation gaps complicate the evaluation of the extent and characteristics of this potential decline of wild bee species in Europe.

According to the European Red List of Bees, 1,101 species (57% of the total) are classified as "data deficient"(Nieto et al. 2014). This lack of scientific information makes it difficult to assess vulnerability and extinction risk for individual taxa. This is important as there are considerable differences in pollination effectiveness and floral specialization (oligolecty vs. polylecty) among genera and species are observed (Dogterom et al. 1998; Cane and Sipes 2006). The misinformation regarding regional wild bee species diversity and distribution has been accompanied by the persistent decline of traditional professional and amateur taxonomic experts since the mid-20th century, a

situation that threatens the future of conservation efforts (Hopkins et al. 2002; Wägele et al. 2011).

DNA-based approaches and in particular DNA metabarcoding might act as a game changer for wild bee assessments (Taberlet et al. 2012; Piper et al. 2019). This tool can generate and process large quantities of data, providing at the same time a way to distinguish cryptic or hard to identify sister species. Moreover, it allows identifying complicated cases such as when facing juveniles or incomplete organisms. Despite its overall potential, the development and applicability of DNA metabarcoding approaches is still a work in progress that has to be evaluated on a case-by-case and regional basis (Leese et al. 2018; van der Loos and Nijland 2020). Metabarcoding datasets are sensitive to multiple factors which can introduce false negative (e.g. gaps in reference libraries, variable primer efficiencies among target taxa, i.e. primer-bias, and variable biomass among specimens and species compromising detection rates) or false positive results (e.g. cross-contamination, tag-switching and a priori identification errors in barcode libraries) (Clarke et al. 2014; Elbrecht and Leese 2015; Weigand et al. 2019; Zinger et al. 2019).

In this study, we tested the suitability of a DNA metabarcoding approach customized for the assessment of the wild bee biodiversity of the Grand Duchy of Luxembourg. Early metabarcoding data already indicate a potential benefit of this approach for assessing Central European wild bees (Gueuning et al. 2019, for Switzerland), but its methodological performance has yet to be evaluated for the regional fauna and for different primer pairs separately. By 2021, 349 wild bee species had been described as present in Luxembourg (Cantú-Salazar et al. 2021; Herrera-Mesías and Weigand 2021), a number that is expected to increase over the next years as a result of increased sampling efforts to develop pollinator monitoring programs. Compared to adjacent countries, this amount is similar to the number of species registered in Belgium (398 species) and The Netherlands (366 species), about one third of the species described from France (949 species) and more than half of the species of Germany (over 550 species) (Westrich et al. 2011; Rasmont et al. 2017; Schneider 2018; Vereecken 2018).

Our central aim here is to propose an effective DNA metabarcoding approach, which ultimately can provide robust data on the wild bee species diversity and distribution in Luxembourg, while preserving bulk samples of wild bees as vouchers. Although this at first glance contradicts the often-highlighted "time-and-cost" benefits of DNA metabarcoding approaches, this strategy will enable subsequent morphological investigations in case of peculiar or doubtful findings, and allows the integration of selected specimens in the reference collection of the National Museum of Natural History Luxembourg (MNHNL).

From a technical point of view, the following methodological aspects were evaluated:

a. Completeness of the barcode reference library of Luxembourgish wild bees
   The commonly used Cytochrome C Oxidase Subunit I (COI) barcode region has a high species discrimination power in Hymenoptera (Smith et al. 2008) and a barcoding library for Central European wild bee species has been available for some time (Schmidt et al. 2015, based on the fauna of Germany). Moreover, new wild bee sequences are being uploaded to the Barcode of Life Data system

(BOLD; Ratnasingham and Hebert 2007) on a regular basis, providing a constantly growing reference library. We thus evaluated the proportion of all regionally cataloged bees having available barcode fragments in general, and more specifically, for three widely applied metabarcoding primer pairs in the study of insects.

b. Effect of primer bias on detectability (*in silico* evaluation)
Numerous primers targeting the COI region have been designed for or applied in metabarcoding studies of insects (e.g. Brandon-Mong et al. 2015; Marquina et al. 2019; Piñol et al. 2019). We compared the *in silico* performance of three promising metabarcoding primer pairs from the literature to select appropriate combinations for wild bees.

c. Effect of biomass bias on detectability
Wild bees are a phenotypically diverse pollinator group with considerable interspecific variation in body size (Michener 2007). As such, wild bee bulk samples might be susceptible to detection biases due to differences in individual biomass. This has been shown in metabarcoding pipelines of wild bees, where strong correlations between read numbers and estimated biomass have been found (Gueuning et al. 2019). Since passive sampling strategies such as pan and malaise trapping are commonly used to collect wild bees despite their body size variations, it is not unlikely that biomass-rich and biomass-low specimens get mixed in bulk samples, potentially obscuring the identification of smaller specimens in a parallel analysis. Adding to this, the effects of primer bias and biomass bias can be synergistic or antagonistic for individual species. We thus examined the effect of biomass in the detection capacity of our metabarcoding approach by using mock communities of known composition under different treatments.

Thus, for this study, *in silico*, and *in vitro* approaches were combined to evaluate the sensitivity of a customized metabarcoding pipeline targeting regional wild bee species to common potential error sources: reference library completeness, primer and biomass-related bias. We compared the in silico performance of three popular metabarcoding primer pairs from the literature and then tested these expectations in the laboratory using mock communities. With this strategy, we aim to determine the best candidate for regional wild bee metabarcoding and to evaluate the suitability of the proposed workflow as a potential identification tool to be used in national conservation initiatives in Luxembourg.

Materials and methods

Barcode reference library coverage analyses

Barcode coverage analyses were performed for three different metabarcoding primer pairs: BF2/BR2, mlCOIintF/HCO2198 and LepF1/MLepF1-Rev (Table 1). The selected primer combinations correspond to primer pairs from the literature that have been previously tested for metabarcoding arthropod samples, showing promising

Table 1. Overview of primer pairs evaluated in the *in silico* analysis.

Primer name	Orientation	Fragment length (bp)	Sequence (5'-3')	Reference
mlCOIintF	forward	313	GGWACWGGWTGAACWGTWTAYCCYCC	Leray et al. (2013)
HCO2198	reverse		TAAACTTCAGGGTGACCAAAAAATCA	Folmer et al. (1994)
BF2	forward	421	GCHCCHGAYATRGCHTTYCC	Elbrecht and Leese (2017b)
BR2	reverse		TCDGGRTGNCCRAARAAYCA	Elbrecht and Leese (2017b)
LepF1	forward	218	ATTCAACCAATCATAAAGATATTGG	Hebert et al. (2004)
MLepF1-Rev	reverse		CGTGGAAAWGCTATATCWGGTG	Brandon-Mong et al. (2015)
LCO1490	forward	658	GGTCAACAAATCATAAAGATATTGG	Folmer et al. (1994)
HCO2198	reverse		TAAACTTCAGGGTGACCAAAAAATCA	Folmer et al. (1994)

amplification success rates for insect taxa, in particular for hymenopterans and bees (Hebert et al. 2004; Brandon-Mong et al. 2015; Elbrecht and Leese 2017a; Gueuning et al. 2019). All pairs included at least one degenerate primer (either forward or reverse), with the BF2/BR2 pair presenting the highest combined primer degeneracy.

Since the DNA barcode sequences deposited in the BOLD reference library may cover different regions of the COI gene, the coverage of the specific amplicon of each primer pair was individually checked for the local wild bee fauna. Additionally, the coverage of the prominent COI Folmer region (i.e. the traditional 658 bp long DNA barcode fragment used for animal barcoding) was also checked.

For this purpose, the R package PrimerMiner version 0.18 (Elbrecht and Leese 2017a) was executed under R version 3.6.2 (R Core Team 2019) to batch download COI sequence data from BOLD (data retrieved on 26-05-2021, minimal barcode length of 196 bp) for the considered 349 Luxembourgish morphospecies, thereby generating the overall barcode coverage report of the public library. As a first step, all available barcode fragments corresponding to the target wild bee species based on their BOLD record details were downloaded and counted. Species were classified into three categories depending on the number of available barcodes (no barcode; 1–4 barcodes; ≥ 5 barcodes).

From this original dataset, identical sequences were automatically reduced to singletons and clustered into Molecular Operational Taxonomic Units (MOTUs) based on a 3% sequence similarity threshold to reduce the bias introduced by unequal representation of sequences in the database (Elbrecht and Leese 2017a).

To further validate the correspondence of each MOTU consensus sequence with the taxonomic data from the BOLD record details of the original barcodes, the resulting fasta files were compared against BOLD Systems using BOLDigger (Buchner and Leese 2020). Taxonomic identification was conducted based on the following sequence similarity thresholds: over 85% of match for identification to the level of order, 90% to family, 95% to genus and 98% to species.

Even if only sequences uploaded under regional wild bee species binomial names were downloaded, the best BOLD match for some MOTU consensus sequences was

not a wild bee present in Luxembourg. For example, OTU 416 consists of a single sequence (BOLD Process ID NOBEE085-09). Even if this sequence was downloaded as *Lasioglossum fratellum*, the available barcode matches the mosquito *Aedes canadensis*. Problematic MOTUs such as this one were deleted from the dataset, keeping only consensus sequences matching Luxembourgish wild bees up to species level.

These validated MOTU consensus sequences representing each target species were subsequently analyzed thus to determine their COI coverage for all three metabarcoding primer-pairs (Suppl. material 1: table a). Finally, the congruence of species delimitation (taxonomic splitting or lumping) was evaluated for each of the 349 morphospecies by tracing back the accession numbers of the original batch downloaded sequences assigned to each Linnaean species across the MOTUs generated in the previous step (Suppl. material 1: tables b, c).

In silico primer evaluation

The PrimerMiner package was used to perform an *in silico* evaluation of the amplification efficiency of each metabarcoding primer based on the dataset of validated MOTU consensus sequences. Scores for primer-template mismatches were assigned based on position and mismatch type under default settings, using the tables included in the package. These scores were summed up to calculate individual penalty scores for each primer or primer pair (Elbrecht and Leese 2017a).

Consensus sequences were visualized with Mesquite v3.6 (Maddison and Maddison 2019) and aligned against a reference "backbone" sequence, obtained by combining wild bee mitochondrial genomes from GenBank into a consensus sequence using MAFFT v7.450 (Katoh et al. 2002). A penalty score of 100 was defined as the threshold (value taken from the package documentation) to determine if a particular primer or primer pair was suitable for the amplification of a specific taxonomic unit. Primers with penalty scores above this threshold were considered inappropriate for metabarcoding. The Folmer primers (LCO1490 and HCO2198) were included in the analysis for comparison. Amplification success rates for each primer were calculated based only on the scores of MOTU consensus sequences with complete sequence data in their respective primer binding sites. Therefore, analogous calculations for the primer pair could only be achieved when complete sequence data was available for both the forward and reverse primer.

To compare the overall performance of the primers across different taxonomic groups, mean penalty scores were calculated by averaging the penalty scores of all the MOTUs within each wild bee genus. Mean values were transformed with a Tukey's Ladder of Powers transformation ($\lambda = 0.375$) to correct for skewness caused due to the presence of outliers (Suppl. material 2). The R packages car (Fox et al. 2016) and rcompanion (Mangiafico and Mangiafico 2017) were used to calculate Shapiro-Wilk and Levene's tests to account for the assumptions of normality and homocedasticity.

To determine whether there were significant differences among the primer pairs regarding their mean *in silico* scores across the wild bee genera, the transformed values

were compared with a weighted One-Way ANOVA, using the number of MOTUs in each genus as weights and the primer pairs as the grouping variable. A Tukey Honest Significant Differences (Tukey's HSD) test was used to calculate pairwise-comparisons between the mean scores of the primer pairs. Both tests were performed in the R package "stats".

Sampling, identification and laboratory processing of specimens

Wild bees were sorted from collections taken in spring and summer 2019 across Luxembourg and the nearby Federal State of Rhineland-Palatinate (Germany) using sweep netting, opportunistic sampling of dead specimens and different kinds of passive trapping (pan traps, vane traps and malaise traps). Wild bee specimens were morphologically identified to the level of species or genus using the taxonomic keys of Amiet et al. (1999, 2001, 2004, 2007), Scheuchl (2000) and Falk (2015).

In the case of the wild bees from Luxembourg, samples were collected over several days using traps filled with 80% propylene glycol and soap or soapy water (Weigand et al. 2021). Individually collected specimens were immediately stored in 96% ethanol. Bees were separated from by-catch in the laboratory.

Wild bees collected in Rhineland-Palatinate were stored in 80% ethanol after sampling, pinned by the end of the field season and kept dry in a drawer. Except for *Ceratina chalybea*, all the specimens from Germany corresponded to species present in Luxembourg (Suppl. material 3: table a). This species was added since no other regional *Ceratina* specimen was available. In total, 32 specimens were selected from the 2019 field work campaigns for the mock community setup. Additionally, a single pinned specimen from the reference collection of the MNHNL and 10 dried specimens opportunistically collected in 2018 and 2019 (found dead in the field) were added as well. This experimental design intended to include representatives from as many of the available genera found in the country as possible in the mock communities, considering only one specimen per species, so that sequencing reads could be easily traced back to a single specimen.

For validation purposes, as well as to check the general suitability of the obtained tissue material for molecular analysis, all specimens were individually Sanger-sequenced using the Folmer primer pair LCO1490/HCO2198. All DNA extractions were performed by grinding a single mid-leg of each specimen in a Retsch TissueLyser Mixer Mill model MM200 using 3 mm beads made of either plastic (41 specimens) or metal (2 specimens), as described in the laboratory protocols of Weigand and Herrera-Mesías (2020). Polymerase chain reaction (PCR), PCR purification and Sanger sequencing were also done according to the protocols described in this publication. In the case of specimens for which it was not possible to get reliable individual barcode sequences using the Folmer primers, molecular identifications were obtained using the LepF1/MlepF1-Rev or the BF2/BR2 primer pairs, using the same PCR thermal profile. Molecular species identification was performed by comparing the obtained sequences against BOLD Systems.

The final assortment of specimens used for the mock community design included 43 adult females. Of them, 28 specimens (25 fresh and 3 dry) were used for "concentration adjustment" mock communities and 14 specimens (7 fresh and 7 dry) for "regular" bulk extraction mock communities, plus a single dry specimen that was used for both treatments (Suppl. material 3: table a). Specimens were classified into three categories based on the overall pre-PCR DNA concentration. Specimens with a concentration below the first quartile of the group were grouped in the small ("S") category, specimens with concentrations between the first and the third quartile were included in the medium ("M") category and specimens with concentrations above the third quartile were assigned to the large ("L") category (Suppl. material 3: table b). DNA concentrations were quantified using a Microvolume Spectrophotometer Trinean Xpose with the A260 dsDNA setting.

Mock community design

To study the *in vitro* effect of primer bias and the impact of biomass differences on the metabarcoding pipeline detection capacity, three experimental set-ups ("mock communities") were designed (Suppl. material 3: table a).

The first mock community (homogeneous, HOMO) was arranged by pooling 10 ng of DNA of each species. In two cases (*Hylaeus nigritus* and *Lasioglossum morio*), just 5 to 6 ng were added due to a lack of further tissue material. This roughly homogeneous treatment provides a theoretically biomass-related bias free scenario, in which differences in the detection rates are more likely to be caused by factors such as primer bias and the stochasticity of the PCR process.

The second mock community (heterogeneous, HETE) was assembled by pooling 1 ul of variable DNA concentration obtained from a single mid-leg from each specimen. This setup provides information regarding the detection limits of the pipeline when DNA from one leg per specimen is analysed and as such, how species are recovered by metabarcoding when the amount of species-specific template DNA is unequal in the PCR. In this case, false negatives are expected to be caused by biomass-related bias under unaltered conditions.

The third mock community (gradient, GRAD) uses the same specimens as in the two previous mock communities, but modifying their concentrations based on the concentration categories previously assigned to each specimen. The DNA of the bees from the "S" category was diluted with buffer in a proportion of 1:100. Six bees from the "M" category were randomly selected and their isolated DNA was diluted to approach concentrations similar to the bees from the "S" category. The bees belonging to the "L" category were not modified. This treatment creates a gradient of DNA concentrations to test the effectiveness of the pipeline under variable DNA concentrations, indicating the impact of the biomass-related bias under more extreme conditions.

Additionally, two "regular mock communities" (RmockA and RmockB) were analyzed for reference purposes (Suppl. material 3). In this case, legs from eight specimens per sample, without repetition of species within the sample, were combined to produce bulk samples from which DNA was isolated.

Metabarcoding PCR, replication strategy, library preparation and sequencing

Three PCR replicates for each mock community (i.e. of HETE, HOMO, GRAD, RmockA and B) were set-up and sequenced. The 16 samples (15 mock community replicates plus a negative control) were amplified using the primer set showing the best performance in the *in silico* evaluation. A two-step PCR protocol was used. The first PCR reaction consisted of 1× Master Mix (GoTaq G2 Hot Start Colorless), 0.5 uM of each primer, 25 ng of DNA and Nuclease-Free H_2O to a final volume of 25 ul. For the second PCR, 1 ul of the amplicon (without cleanup) was used as template and the amount of reactives was modified to a final volume of 50 ul. Both PCRs were run on an Eppendorf Mastercycler nexus eco thermocycler using thermal profiles based on the ones described in Elbrecht and Steinke (2019). The first PCR started with an initial denaturation step at 94 °C for 5 minutes, followed by 34 cycles of denaturation for 30 seconds at 94 °C with annealing for 30 seconds at 50 °C and extension at 65 °C for 50 seconds; and a final extension for 5 minutes at 65 °C. The program for the second PCR followed the same steps, but with 19 cycles instead and an extension time of 2 minutes. The tag combination used for the second PCR are described in Elbrecht and Leese (2017b) (Suppl. material 4). PCR success was verified by electrophoresis and the products were purified with a NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel™). The DNA concentrations of the purified products were measured for equimolarly pooling into the final library (40 ul, 64.3 ng/ul). The cleaned library was sequenced on one lane of an Illumina MiSeq System (2 × 250 bp) at the Luxembourg Centre for Systems Biomedicine (Belval, Luxembourg).

Quality filtering, MOTU clustering and taxonomic assignment

Dereplication of the samples in the same sequencing run based on inline tag combinations was done using scripts ("Demultiplexer") developed by the Aquatic Ecosystem Research Group of the University of Duisburg-Essen. Reads that were unmatched after this module were mapped against PhiX to check for the presence of virus genome. The demultiplexed data was further processed using the JAMP ("Just Another Metabarcoding Pipeline") R package (https://github.com/VascoElbrecht/JAMP). This package consists of a modular metabarcoding pipeline that provides extended quality filtering options and automatically generated summary statistics, integrating different functions from external programs to produce the output of the different steps (Elbrecht et al. 2018). JAMP 0.67 was run using R version 3.6.2 (R Core Team 2019), relying on Usearch v11.0.667 (Edgar 2010), Vsearch v2.14.2 (Rognes et al. 2016) and Cutadapt 2.8 (Martin 2011). Settings were adjusted so that 25% mismatches were allowed to overlap. Any read that did not match the expected length of the BF2/BR2 amplicon (421bp ± 10bp) was removed. After MOTU clustering based on 3% sequence similarity, a default 0.01% abundance filter was applied twice (i.e. first based on the overall dataset and second based on the results of each individual sample) to the initial MOTU table to produce the final dataset.

Taxonomic sorting was performed by comparing the resulting MOTU fasta files against sequences stored in BOLD Systems using BOLDigger. The same thresholds for taxonomic identification used in the *in silico* evaluation were used here. The resulting data were pruned using TaxonTableTools (Macher et al. 2021) to remove all non-Hymenoptera MOTUs, as well as Hymenoptera MOTUs present in only one out of the three PCR replicates.

The detection capacity of the metabarcoding pipeline was evaluated based on the percentage of intentionally pooled species retrieved from each treatment ("detection rates"). Kruskal-Wallis and Wilcoxon rank sum tests (both default R package "stats 3.6.2") were used to determine whether there were significant differences among the HETE, HOMO and GRAD mock communities regarding the average read numbers per species obtained after combining the sequencing results of all replicates.

Results

Barcode coverage

Of the 349 wild bee species evaluated, 96.84% presented at least one COI barcode sequence available in the BOLD Systems public library (Fig. 1); 84.81% of all morphospecies were represented by at least five barcodes and 12.03% by one to four barcodes. Only eleven species (3.15%) had not barcodes in the database.

The 7,317 de-replicated sequences considered (i.e. after removing identical sequences from the set of 11,810 downloaded sequences) were clustered into 558 MOTUs. From them, the consensus sequences of 433 were included in the final dataset based on the combined assessment of their 20 top matches using BOLDigger, and supporting their identification as local wild bee taxa.

Barcode coverage of the regions targeted by the three considered metabarcoding primer pairs presented little variation (Fig. 2). The mlCOIintF/HCO2198 target fragment was completely covered for 77.94% of the morphospecies, partially covered for 9.74% and it was missing for 12.32%. LepF1/MLepF1-rev presented a complete coverage in 77.65% of morphospecies, a partial coverage for 10.60% and for 11.75% the target region was missing in BOLD. The BF2/BR2 target fragment was complete for 68.48% of the species, partially covered for 20.34% and missing in 11.18%. Full-length (complete) barcode coverage for the traditional animal barcoding Folmer primer pair LCO1490/HCO2198 target region was the lowest (59.89%).

Species delimitation congruence

Only 39.05% (132/338) of the morphospecies considered in the final dataset fulfilled the expected correspondence of one MOTU per Linnaean species (Suppl. material 7). In all the other cases, some sort of incongruence was observed (Suppl. material 1: table b).

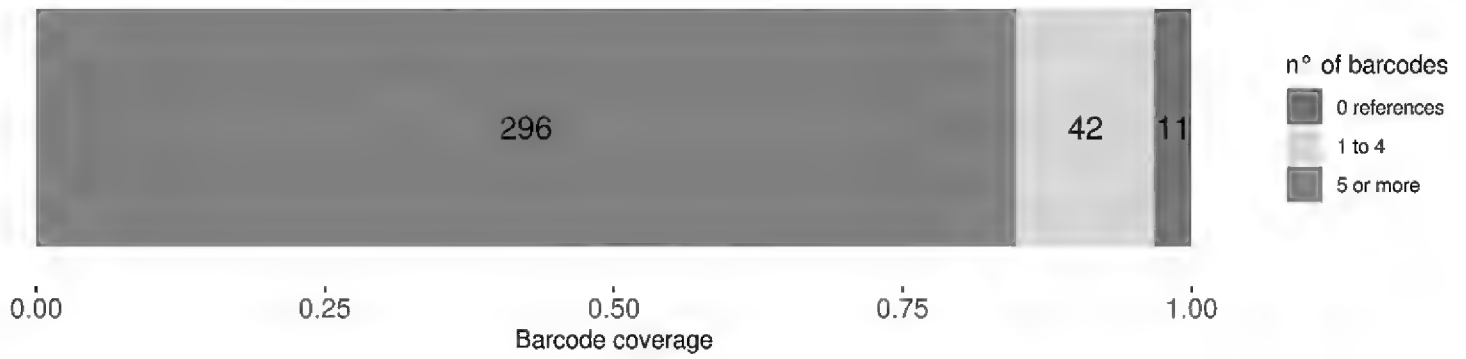

Figure 1. Overall COI barcode coverage in the BOLD public library for the 349 Luxembourgish wild bee species considered.

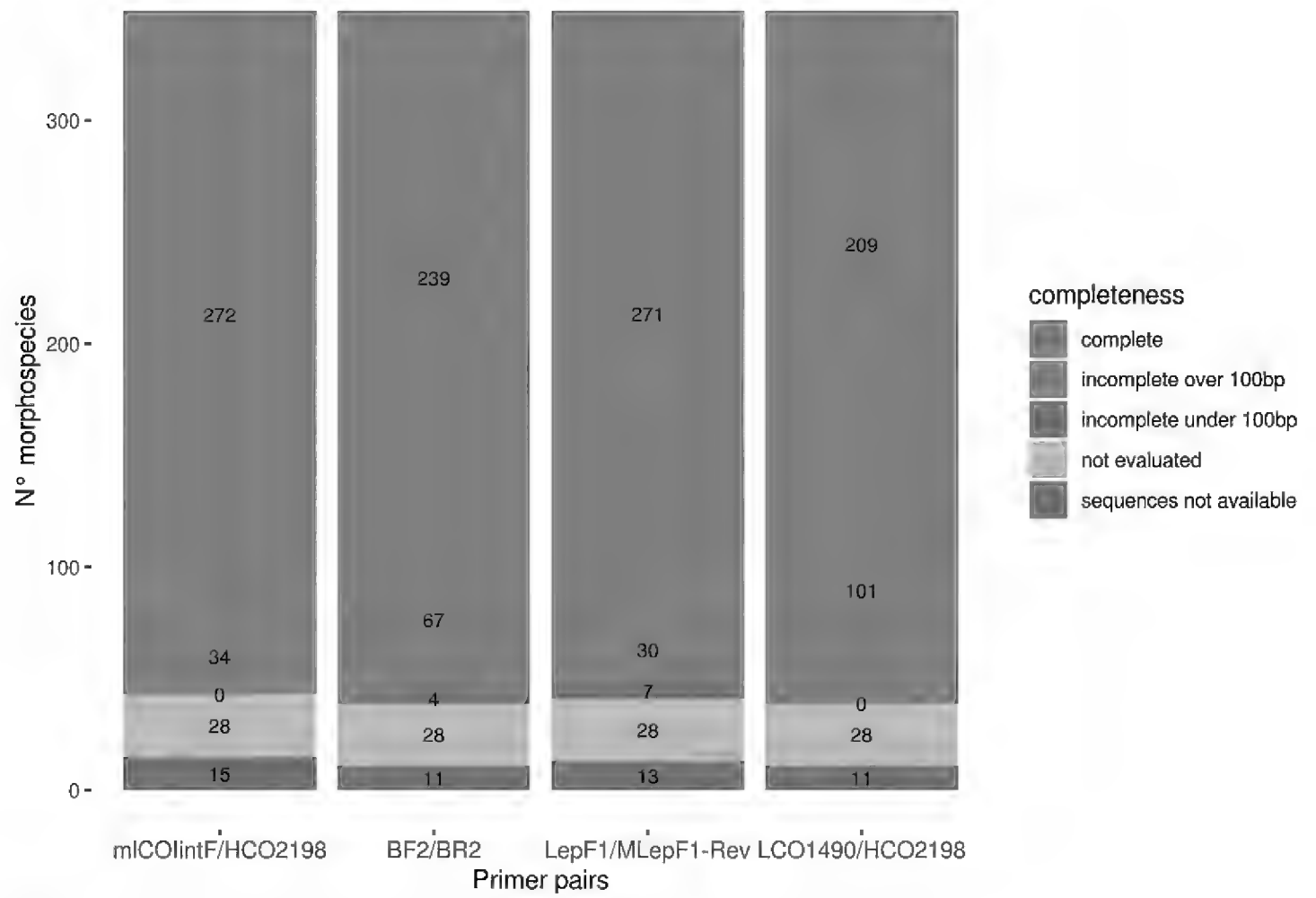

Figure 2. Coverage of (meta)barcoding primer pairs considered in this study for the targeted 349 Luxembourgish wild bee species.

The clustering process either split sequences belonging to one species into multiple MOTUs (29.29% of the cases; mostly into two MOTU) or lumped sequences from different morphospecies into a single MOTU (8.88% of the cases). While lumping might lead to false negatives or unresolved metabarcoding results, splitting of a nominal species into multiple MOTUs generally does not bias metabarcoding outcomes. Finally, in 22.78% of the cases, the 3% clustering threshold split the sequences of a Linnaean species and then lumped them with sequences corresponding to another morphospecies. This process created MOTUs that combined sequences from several species.

In silico primer evaluation

The results of the *in silico* analysis of the metabarcoding primers were first sorted by genera to assess their performances across different taxonomic groups of interest. When all MOTUs are considered, the expected amplification success rates of the individual primers varied across 25 wild bee genera. However, in the majority of the cases the combined outcomes of the metabarcoding primer pairs were higher or equal to the ones of the standard Folmer barcoding primers (Suppl. material 8). Exceptions to this were the genera *Sphecodes*, *Osmia*, *Hoplitis*, *Halictus*, *Megachile*, *Chelostoma*, *Colletes* and *Dasypoda*, for which the Folmer primers outperformed one or more metabarcoding primer pairs.

The BF2/BR2 primer pair had the highest mean *in silico* amplification success rate (86.52% of the species with binding site sequence data were expected to correctly amplify), while LepF1/MLepF1-Rev (16.88% of the species) and LCO1490/HCO2198 (17.65% of the species) had the lowest success rates. The primer pair mlCOIintF/HCO2198 showed an intermediate *in silico* performance (amplification is expected successful for 37.50% of the species). The expected amplification success rates of the primer pairs mlCOIintF/HCO2198 and BF2/BR2 were identical for 48.57% of the wild bee genera considered. However, BF2/BR2 consistently outperforms mlCOIintF/HCO2198 in 83.33% of the remaining cases, while mlCOIintF/HCO2198 only shows higher amplification success rates than BF2/BR2 in three genera: *Nomada*, *Heriades* and *Melitta*.

Regarding the average penalty scores obtained from all the MOTUs within each wild bee genus, the transformed scores for BF2/BR2 were within the accepted values of amplification success, with the exception of the mean penalty scores of *Anthophora*, *Eucera*, *Halictus*, *Melitta* and *Nomada* (Fig. 3). In contrast, the average penalty scores of only nine genera were below the threshold for mlCOIintF/HCO2198 and of only two genera for LepF1/MLepF1-Rev. Moreover, the BF2/BR2 mean score calculated from all genus average penalty scores was the only one below the threshold. The results of the weighted One-Way ANOVA indicated that there was a statistically-significant difference in the genera average transformed penalty scores by metabarcoding primer pair ($f(2) = 42.98$, $p < 0.001$). The Tukey's HSD test indicated that the differences were statistically significant for all primer comparisons ($p < 0.001$ for all pairwise comparisons). Data is normally distributed and homocedastic at a 95% level of confidence (Shapiro-Wilk test: $W = 0.98$, $p = 0.241$; Levene's Test: $F(2) = 2.47$, $p = 0.092$).

The results of the *in silico* analysis vary slightly when only the MOTU with the best score for a distinctive morphospecies (in the case of "multi-MOTU" species) is considered as an outcome (Suppl. material 9). Under this assumption, the metabarcoding primer pair with the highest amplification success rate is BF2/BR2 (87.05%), followed by mlCOIintF/HCO2198 (36.09%) and finally LepF1/MLepF1-Rev (17.44%).

Overall, multi-MOTU morphospecies presented congruent results for the same primer pair, despite variable penalty scores for each of their MOTUs. The exception to this were four species (*A. plumipes*, *B. terrestris*, *S. albilabris*, and *S. geoffrellus*), which presented MOTUs with scores both above and below the threshold for one or more primer combinations. *Bombus terrestris* presented discrepancies for all primer pairs but

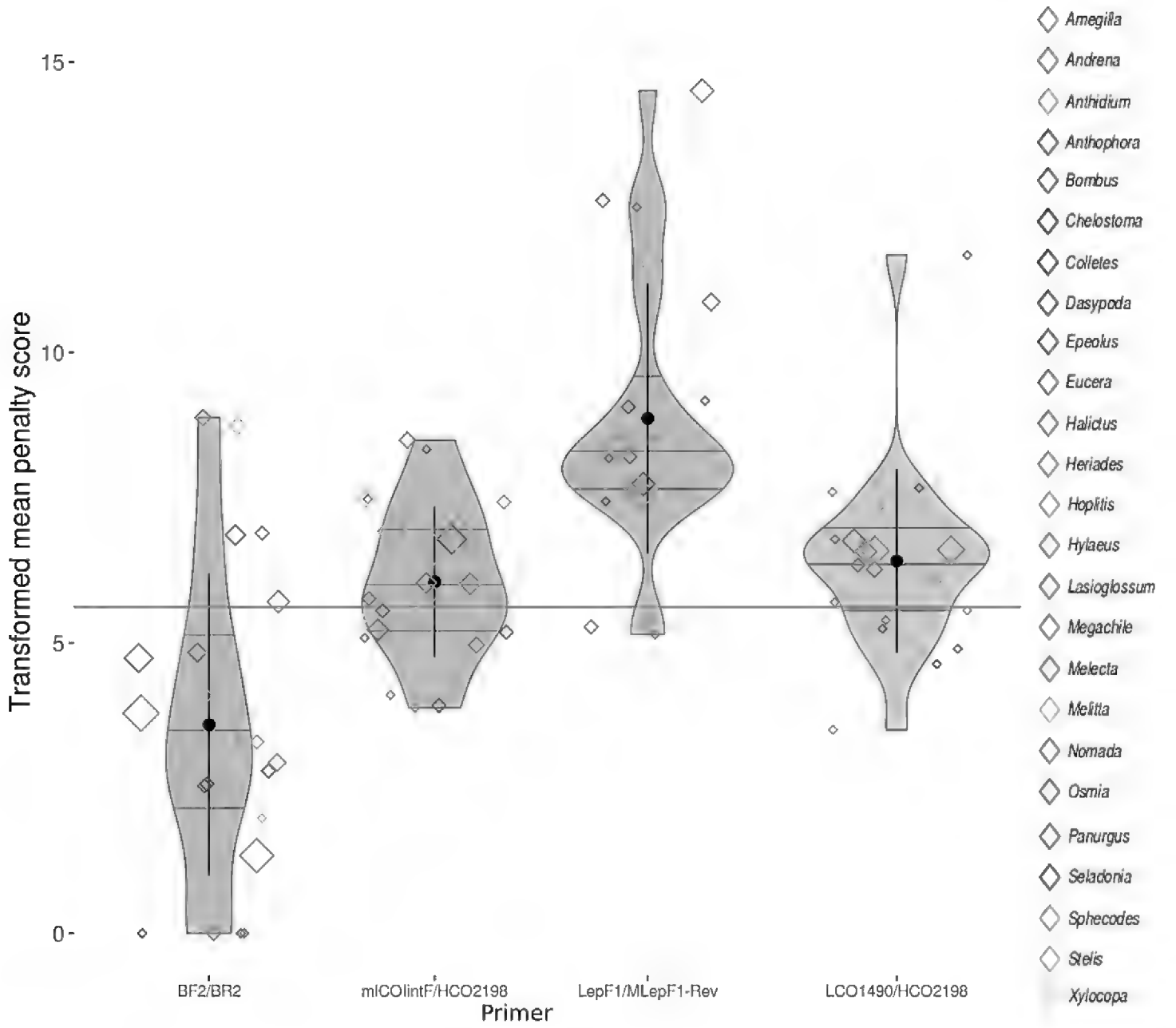

Figure 3. Distribution of transformed mean penalty scores by primer pair and genus. The sizes of the diamonds that represent the mean of the genera are proportional to the number of MOTUs in each genus. The overall mean value and standard deviations for each primer pair are shown in black. Means with a penalty above the red line (penalty score of 100) are considered *in silico* performance failures. The LCO1490/HCO2198 primer pair is indicated for reference purposes only, and not included in the statistical analysis.

LepF1/MLepF1-Rev. Two species (*A. plumipes*, and *S. geoffrellus*) showed discrepancies for mlCOIintF/HCO2198 and BF2/BR2. Finally, *S. albilabris* only presented discrepancies for mlCOIintF/HCO2198.

Bioinformatic analysis of mock communities and detection rates

A total of 6,902,568 high quality reads from the original 11,701,736 read pairs remained after trimming and quality filtering (Short Read Archive bioproject number PRJNA867321). The percentage corresponding to PhiX found in the unassigned reads (64% of the 2,251,231 reads in "no match") was in agreement with the procedures of the sequencing center. From the original 328 MOTUs generated, 118 MOTUs

remained after the 0.01% abundance filters (Suppl. material 5). 1,126 chimeras were discarded during clustering. The sample presented a moderate level of resolution, with 70% of the MOTUs identified at least up to the level of genus and 52% to the level of species.

For species detection rate assessment within mock communities, 53 Hymenoptera MOTUs – identified to species level and present in at least two replicates – were considered (Suppl. material 6). The detection rate of input species was 97% for both the HETE and HOMO mock communities, but only 72% in the GRAD mock community. The single missing species in the HETE sample corresponded to a "S" category species (*L. morio*) that was only found in the first replicate of the sample with 53 reads, while the missing bee in the HOMO sample corresponded to a "L" category museum specimen (*T. byssina*), whose DNA was potentially already degraded. All the missing species in the GRAD mock community belonged to the "S" category (*H. langobardicus*, *C. afra*, *L. morio*, *L. nitidulum*, *E. alticincta*, *H. tumulorum*, *L. laticeps* and *H. nigritus*), except for *T. byssina*, which was found only in the first replicate of the sample with 110 reads. Bee specimens in the "M" category were detected in all three set-ups, even when diluted in a proportion of 1:100 (GRAD mock community). Both regular set-ups (RmockA and RmockB) had a detection rate of input species of 100%. However, a false positive (*Halictus confusus*) was found in both, likely due to pre-PCR contamination.

In the main three experimental set-ups (HETE, HOMO, GRAD), sequence reads of *Andrena cineraria* dominated the results, with over 30% of the average reads in all three mock communities and replicates (Fig. 4). In the HETE and GRAD mock communities, *Bombus lapidarius* and *Dasypoda hirtipes* were both highly represented (*B. lapidarius*: 25% to 21% of the reads, *D. hirtipes*: 15% to 13%), but not in the HOMO mock community (*B. lapidarius*: 11%, *D. hirtipes*: 6%). The number of reads corresponding to the wild bee with the highest biomass among all specimens pooled in the mock communities (*Xylocopa violacea*) was neither particularly high in the HETE nor in the GRAD treatment, and it has considerably less reads than *A. cineraria*.

The results of the Kruskal-Wallis rank sum test and Wilcoxon rank sum test indicate the presence of significant differences in average read numbers per species only between the GRAD and the HOMO mock community at a 95% confidence level (Kruskal-Wallis $x^2(2) = 8.12$, $p = 0.017$; Wilcoxon rank sum test with Bonferroni correction $p < 0.05$ only for GRADxHOMO comparison). No significant differences in average read numbers per species were found between the HETE mock community and the two other treatments. Data is not normally distributed but homoscedastic (Shapiro-Wilk test: $W = 0.46$, $p = 2.597e^{-16}$, Levene's Test: $F(2) = 0.004$, $p = 0.96$). Also, it is important to mention that nine morphospecies were represented by multiple MOTUs in the metabarcoding results, despite only one specimen being pooled in the mock community mixtures (Suppl. material 6): *A. carantonica*, *B. lapidarius*, *B. terrestris*, *C. afra*, *C. cunicularius*, *D. hirtipes*, *E. interrupta*, *H. tumulorum* and *O. bicornis*.

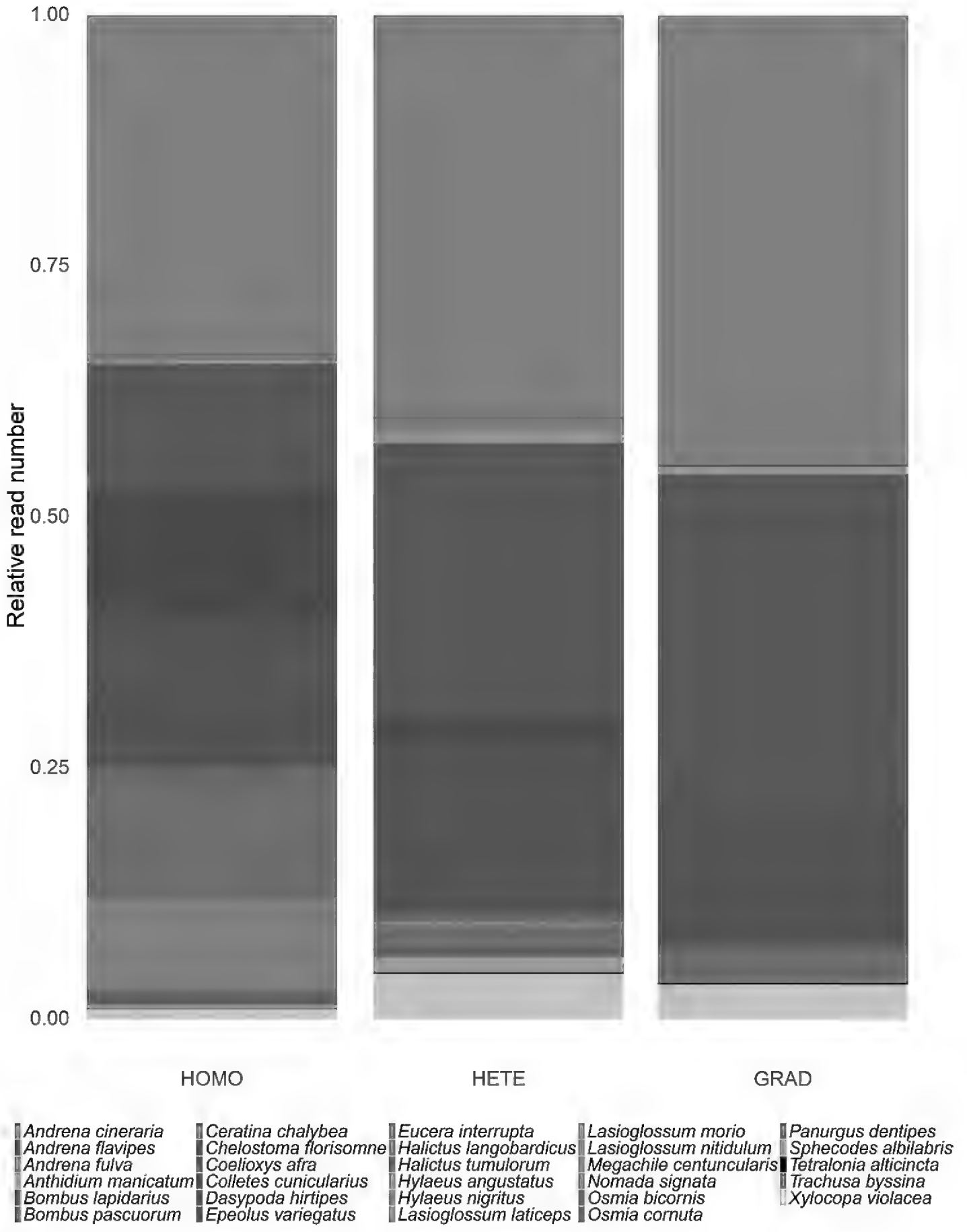

Figure 4. Proportion of sequencing reads per species in each mock community. Each community was assembled from the DNA of the same 29 specimens (25 fresh ones and 4 dry ones), all of them belonging to different species; the HOMO community was based on equimolar pools of the individual DNA extractions; the HETE community was assembled by pooling 1 ul of isolated DNA from a single leg from each specimen and the GRAD community was made by modifying the original concentrations of each species based on their concentration categories in order to exaggerate existing biomass differences. Relative read numbers were obtained by averaging absolute read numbers from all three replicates of each mock community and then correcting by the total number of reads in each treatment. A significant difference was found only between GRAD and HOMO (see text for details).

Non-Hymenoptera metazoan MOTUs found in the samples, such as *Parus major* and *Nephrotoma appendiculata*, are likely the result of contamination with organisms present in the field. The fungi and plant DNA found in the samples are also likely to be due to carry over from the field or rather contaminations with materials from other research groups at the laboratory of the MNHNL.

Discussion

Barcode coverage analysis

The availability of reference barcode sequences is a central requirement when evaluating the performance power of a DNA-based identification method for a certain taxonomic group, geographical region or environment (Weigand et al. 2019). The barcode coverage analysis here performed shows that missing barcodes are not a general limitation for DNA metabarcoding analysis of the local wild bee fauna of Luxembourg. In total, 338 morphospecies (97%) were represented by at least one COI barcode sequence in BOLD systems. Among them, 296 species are particularly well covered (>5 sequences available). However, these results only indicate the presence in the database of one or more COI sequences with over 196 bp of length for the target taxa, regardless of their specific location within the COI gene and prominent Folmer region. To properly evaluate the suitability of the proposed metabarcoding approach, the coverages of the selected target fragments had to be evaluated case-by-case. The coverage analyses of the target fragments of the three selected metabarcoding primer pairs shows that a full length fragment can only be expected for 68%–78% of the species, despite the primer pair combination. If species with partial coverage (>100 bp) are added to the primer pair-specific evaluations ("mini-barcodes", Meusnier et al. 2008), over 85% of the wild bees currently described in Luxembourg have a reference in BOLD, regardless of which metabarcoding primer pair is used. Since PrimerMiner is only able to batch download publicly available sequences, the number of BOLD reference sequences might even be slightly higher when comparisons make use of the full database (incl. non-public records). Moreover, a few morphospecies were omitted by the automated metabarcoding pipeline due to i) alignment artifacts in Mesquite, ii) clustering of a morphospecies into multiple MOTUs with different taxonomic IDs (this was only the case for *N. striata*), and iii) an extreme difference between the original annotation of a MOTU consensus sequence and its BOLDigger re-identification. The latter problem may be partly related to limitations of the PrimerMiner algorithm generating the consensus sequences, which was primarily developed and is frequently used for MOTU consensus sequence construction at the taxonomic level of order and/or family.

Congruence of MOTUs with morphospecies

With the applied clustering threshold value of 3% sequence similarity, only 39% (132/338) of the evaluated wild bee species met the expectation of one MOTU per

morphospecies based on a Linnaean species delimitation concept. In 29% of the cases, a morphospecies split into multiple MOTUs, while in 9% of the cases sequences from multiple morphospecies lumped together. Additionally, 23% of the cases showed a variable combination of both effects, including multiple barcodes merging into mixed species MOTUs. For example, the COI barcodes downloaded for *Andrena bimaculata* split into two MOTUs. Three barcodes clustered together with *Andrena tibialis* in MOTU 203, while the remaining five formed a mixed species MOTU (MOTU 345). These deviations are in agreement with the incongruences described by Creedy et al. (2020) for the wild bee fauna of the United Kingdom. Their phylogenetic analyses suggested that these deviations could be due to closely related taxonomic groups and/or to the geographical range of available DNA barcodes (Creedy et al. 2020). The effect of this latter factor might be avoided by only considering DNA barcodes from local sources (Bergsten et al. 2012), which was not possible in our case but must be aimed at.

It is worth noticing that at least part of the splitting and lumping situation observed here is potentially the result of sequences uploaded under incorrect species annotation into BOLD. Outstanding examples can be found in the DNA barcode material of *Nomada striata*, which split into three MOTUs and then lumped with different morphospecies in each mixed MOTU (MOTU20: *N. ruficornis* and *N. fulvicornis*; MOTU259: *N. alboguttata*; MOTU310: *N. zonata*). Furthermore, the BOLD_BIN ABY7961 of *N. striata* not only includes annotated specimens of this species, but *N. villosa* (4 specimens) and *N. symphyti* (1) -two species so far not reported for Luxembourg (Cantú-Salazar et al. 2021; Herrera-Mesías and Weigand 2021) and hence not considered by us. A single specimen can be found in BOLD_BIN AAF3496, identified as *N. striata* but most likely corresponding to *N. zonata* based on their genetic data. Improved quality control of DNA barcode voucher material and its associated metadata is advisable to reduce potential noise in the database (Weigand et al. 2019). A well-curated regional database for the wild bee fauna of Luxembourg comprising a few but high-quality entries per species might help to overcome similar MOTU annotation problems in the future.

Pipeline evaluation and potential error sources

Even in cases when a reliable reference barcode library is available for the target taxa, primer bias can lead to false negatives and/or reduced detection rates (Elbrecht and Leese 2015). False negative results can also be generated when a low-biomass specimen is analyzed in parallel with high-biomass specimens or in a generally biomass-rich sample (Elbrecht and Leese 2015). Hence, it is of paramount importance to understand the effects of non-equal primer binding (amplification) efficiencies and variable biomass differences for the taxonomic groups under study. Our *in silico* evaluation of three metabarcoding primer pairs consistently identified the BF2/BR2 primer pair as the top performer for local wild bee assessment: over 85% of all MOTUs and morphospecies for which complete binding site sequence data was available are expected to efficiently amplify based on their simulated amplification success rates. However, deviations from these expectations set by the *in silico* analysis can potentially be found in laboratory set-

ups due to several factors. Even if primer-template mismatch has been experimentally shown to have a disproportionate effect over amplification success in mock communities (Piñol et al. 2015; Piñol et al. 2019), other factors such as annealing temperature, PCR cycle number or blocking oligonucleotide concentration can also affect species relative abundance in metabarcoding analyses (Piñol et al. 2015). Interestingly, species from genera predicted to present amplification troubles based on their mean penalty scores (i.e. *Anthophora*, *Eucera*, *Halictus*, *Melitta* and *Nomada*) correctly amplified in our mock communities and were easily detected among the pipeline results. Further laboratory experiments are needed to evaluate the actual amplification efficiency of the BF2/BR2 primer pair in potentially troubling wild bee taxa, thus to adjust expectations and uncover other potential factors affecting metabarcoding results.

COI metabarcoding approaches rely on degenerate primers such as BF2/BR2 to maximize taxon recovery, as this allows matching at variable binding sites and the amplification of as many (target) input sequences as possible (Linhart and Shamir 2002; Elbrecht et al. 2018). However, high degeneracy increases the chances of co-amplifying non-target sequences, potentially loosing specificity (Linhart and Shamir 2002). Even if these non-target sequences (NUMTs, pseudogenes or parasitic/bacterial contaminants) may be bioinformatically filtered out, such procedure can reduce the recovery of target sequences (Elbrecht et al. 2018), affecting the overall detection capacity of the pipeline. Whenever possible, the susceptibility of specific degenerate primer combinations to this bias should be evaluated and taken into consideration for the experimental design, based on the taxa of interest. In the case of BF2/BR2, laboratory validations performed on invertebrate mock communities indicate that the amplification of non-target regions is minimal when this primer pair is used for insect taxa metabarcoding, with less than 0.5% of all resulting sequences deviating from the expected length (Elbrecht and Leese 2017b). However, conclusions regarding this aspect must be drawn carefully, as subsequent studies have also shown that the BF2 primer is also susceptible to primer slippage, which depending on the target taxa analyzed, may result in part of the amplicon sequences to be a few bp longer or shorter than expected (Elbrecht et al. 2018).

In principle, it must be highlighted that a highly degenerate primer pair can generally perform well in an *in silico* analysis, but might mal-perform *in vitro* due to the co-amplification of non-target taxa.

In our study, we tested the predictions of the *in silico* analysis by sequencing five distinct mock communities using our best performing primer pair (i.e. homogeneous, heterogeneous, gradient and two regular mock communities). The final detection rates of input species for the HOMO and HETE mock communities were the same (97%), while the detection rate of the GRAD mock community was considerably lower (72%). The missing species in the HOMO mock community (*T. byssina*, "L" category) likely represents an artifact, considering that a 7-year-old museum sample with unknown initial preservation conditions was used. This hypothesis is supported by the fact that the fresh specimen of *T. byssina* used for bulk extraction in the regular mock communities was found in all replicates. DNA degradation over time in insect museum samples is a well-known phenomenon and models have been developed to characterize

the level of molecular damage (Zimmermann et al. 2008). Therefore, metabarcoding projects working with preserved insect specimens (i.e. confirming the presence of a species from damaged historical samples to complete museum databases) should be aware of potential DNA damage that may bias their results. All missing species in the GRAD mock community correspond to bees from the "S" category. The samples in this category are bees from the genera *Halictus*, *Coelioxys*, *Hylaeus* and *Lasioglossum*. Originally, all of them had an overall pre-PCR DNA concentration between 4.5 and 1.2 ng/ul, but in the GRAD treatment, they were diluted in a proportion 1:100. This artificial concentration is often well below the expected DNA concentration of a full leg after isolation, even of the smallest Central European wild bee specimen. However, particularly specimen-rich bulk samples containing several *Bombus* spp. and honey bees may complicate the detection of a single small-sized bee species if it is represented by just a few specimens (e.g. *Lasioglossum* spp.), due to the magnitude of the difference between their template DNA compared to the total DNA of the sample.

Since a single specimen per species was pooled in our mock communities, the proportion of sequence reads per species should be similar in all experimental set-ups, unless error sources (i.e primer mismatch, biomass bias, etc) were biasing the relative read abundances, favoring some taxonomic groups over others (Braukmann et al. 2019). Therefore, the differences observed in the proportion of sequences from each taxonomic group in the mock community supports that one or more error sources are affecting the results of the pipeline.

In all three main mock communities (HETE, HOMO, GRAD), 30% to 45% of all reads corresponded to *A. cineraria*, a species that has a considerable biomass (pre-PCR DNA concentration: 48.8 ng/ml, dry weight: 31.9 mg) and a very low primer-template mismatch (penalty score: 18.32). In the GRAD and HETE mock communities, biomass-rich species from the "L" category tended to have higher overall read numbers. However, no significant differences were found in detection rates or in read numbers per species among the HOMO and the HETE mock communities. Therefore, there is no evidence suggesting that correcting for biomass differences (e.g. size-sorting) has a significant effect in the general assessment outcome of our wild bee metabarcoding approach, at least under the conditions here proposed. Hence, isolating a single leg from each wild bee specimen should be sufficient for its detection in an average bulk sample under the described sequencing depth. Nevertheless, it is important to acknowledge that challenging bulk sample mixtures consisting of few small-sized taxa and an overabundance of large-sized bees might result in further problems not evaluated in this study.

In summary, the comparison between the results of the HETE and the HOMO mock communities suggest that the differences found in the proportion of read numbers per species are likely due to differential amplification resulting from primer bias. In the case of the GRAD community, the proportion of input species read numbers was not significantly different from the HETE mock community and the overall detection rate was only mildly affected. Overall, these results suggest that primer bias was the principal driver behind the unequal representation of species in the mock communities, with biomass differences only adding to the effect as a secondary factor.

Quantitative estimations from metabarcoding results: Is it possible?

The results found in the HETE mock community suggest a general trend of biomass-rich bee taxa to have higher read numbers. However, it is unlikely that this information can be used to retrieve accurate quantitative results regarding species biomass or specimen abundances. If PCR-based approaches are used in a metabarcoding set-up, the effect of differential amplification efficiency would make extremely difficult to estimate any of these parameters based on the final read numbers (Piñol et al. 2015, 2019; Elbrecht and Leese 2015). Numerical experiments done with computational simulations using insect datasets indicate that the capacity of providing quantitative estimates regarding the composition of the original sample will largely depend on the primer pair used for amplification and on the characteristics of the species community analyzed (Piñol et al. 2019). In the particular case of BF2/BR2, a significant correlation between pre- and post-PCR DNA concentrations has been reported for insect taxa, suggesting that it would be theoretically possible to quantify the initial abundance of each species in a bulk sample using customized equations, given that the species composition and number of primer-template mismatches are known (Piñol et al. 2019). However, the metabarcoding pipeline here developed should only be used for the qualitative assessment of wild bee fauna, at least until this hypothesis is experimentally tested and further data regarding quantitative estimations using the BF2/BR2 primer pair become publicly available.

Multiple MOTUs originating from single specimens

It is noteworthy that multiple MOTUs from the same species were found among the mock community metabarcoding results, despite a single specimen being used for the design. The presence of multiple MOTUs may have been caused by the effect of mitochondrial heteroplasmy or by nuclear copies of mtDNA (numts). The presence of multiple mitochondrial DNA haplotypes coexisting in a single organism remains a potential problem for the use of DNA (meta)barcoding as a molecular taxonomic tool (Rubinoff et al. 2006). Even if maternal mitochondrial DNA inheritance is considered the general rule for eukaryotes, it has been observed that paternal mtDNA transfer can happen during polyspermic fertilization in honeybees, a fraction of which is partially retained in later developmental stages (Meusel and Moritz 1993). In the case of wild bees, high proportions of heteroplasmic species have been described for Hawaiian *Hylaeus* spp. (Magnacca and Brown 2010), indicating that mtDNA heteroplasmy can occur in wild bees and that it might be more common than originally thought. To the best of our knowledge, this is the only study suggesting heteroplasmy in wild bees and further research would be needed to confirm its findings, especially as it can be difficult to distinguish heteroplasmy from the presence of highly similar NUMTS. In our mock communities, three MOTUs (separated by 3% sequence divergence) originated from a single *Dasypoda hirtipes* female, showing a sequence similarity of 100%, 99.49% and 99.45% with their best BOLD matches. This high sequence similarity and the

congruent detection of those MOTUs in all three replicates of every mock community excludes PCR and sequencing errors as the primary source for the anomaly.

Alternatively, these peculiarities in the dataset may be explained by nuclear mitochondrial DNA (NUMT) sequences. NUMTs are the result of non-translated and non-transcribed regions from the mitochondrial DNA transferred to the nuclear genome, which can be amplified if effective primer binding sites are still existing (Cristiano et al. 2012). This causes the amplification of non-functional nuclear copies of COI together with real mitochondrial DNA, producing a mix of copies that will result in several MOTUs originating from the same specimen (Cristiano et al. 2012). Molecular phylogenetic analysis in an extended dataset including the species here described might be useful to search for evidence of potential COI-like NUMTs in the target taxa. Sample contamination as an explanation for this anomaly (e.g. environmental DNA carry over) seems very improbable, as this would have also likely introduced new species and not only inflated the number of MOTUs of species already present in the mock communities.

Further studies should determine the presence and the potential impact of heteroplasmy and NUMTs in the effectiveness of barcoding identification of potentially heteroplasmic wild bees of both sexes, as well as the impact of multiple MOTUs originating from single specimens on diversity estimates.

Conclusion

The *in silico* and *in vitro* analyses highlight the influence of primer bias on the performance of the proposed metabarcoding approach. However, it is possible to reduce its effect by selecting the most suitable primer combinations for the taxa of interest. This can be achieved by comparing the *in silico* amplification efficiency of primer pair candidates and then experimentally testing the capacity of the best performing pairs in the laboratory. Among the metabarcoding primer pairs here evaluated, no combination can be expected to correctly amplify all wild bee taxa and some genera in particular are predicted to be prone to amplification problems, ultimately translating into a higher probability of producing false negatives. Therefore, primers have to be evaluated on a case-by-case basis against the target taxa at hand. Nevertheless, from the combinations available, the highly degenerate primer pair BF2/BR2 provided the best results for our regional wild bee fauna, with over 85% of available MOTUs and morphospecies expected to correctly amplify when this primer pair is used. Our experimental set-ups support these results as over 97% of the species were retrieved from four out of five mock community trials using the metabarcoding approach that incorporates this primer pair.

A deficiency of DNA barcodes in the public reference library BOLD does not seem to be a major error source for the identification of the regional wild bee species using molecular taxonomic tools. In total, 97% of the currently known morphospecies in Luxembourg present at least one barcode in BOLD, and 85% of them can be considered well covered. However, for the ~30% of the taxa whose identification might be

obscured due to lumping with other wild bee species, the definition of potentially diagnostic barcodes or a multi-marker DNA metabarcoding approach (i.e. incorporating nuclear markers) may be considered as alternative strategies to discriminate lumped species. Finally, new sampling campaigns and collection revisions are likely to provide material to fill the few remaining gaps in the database, as well as to produce barcodes originating from regional specimens.

The results of the mock community experiments indicate that the overall output of the metabarcoding pipeline is expected to be robust, despite biomass differences among the wild bee specimens. Even if these biomass differences affect the number of reads per taxonomic group, the detection rates of input species (i.e. taxalists) remained stable, with the exception of the gradient treatment. Biomass-related bias is likely to have a higher impact under more extreme scenarios, where the size difference of the pooled specimens is higher (e.g. in Malaise traps). Moreover, due to the small numbers of specimens included in this analysis, a higher effect of this type of bias in bulk samples combing numerous biomass-rich specimens and few biomass-low ones cannot be ruled out. However, strategies can be used to compensate for this issue under reasonable conditions. In general, processing separately the fraction of smallest wild bee specimens in a sample should provide an appropriate countermeasure to avoid false negative results due to biomass differences, especially for genera with negative primer bias (e.g. *Nomada* spp.). Moreover, as the proposed metabarcoding pipeline only uses one leg for bulk extraction, the voucher specimens can be traced back for complementary analysis with Sanger sequencing or traditional morphotaxonomy, thus to provide identifications validated by multiple approaches.

Even if only a few specimens were used here to set up the mock community trial, the layout of the metabarcoding pipeline in this study can be used to analyze much larger samples. Sequencing costs for a HTS run on an Illumina platform remain stable independently of how many individuals are included in each bulk sample, and the BF2/BR2 tagging primer combinations allow the tagging of up to 288 samples within the same run (Elbrecht and Steinke 2019). Therefore, the current metabarcoding approach can potentially be used to analyze hundreds of bulk samples containing several dozens of wild bees on a single run without incurring in substantial modifications to the workflow or significantly higher costs.

Overall, our customized metabarcoding pipeline represents a promising alternative taxonomic identification tool to analyze large numbers of wild bees in the context of local conservation biology initiatives. As such, the further improvement of this technique would benefit projects dealing with many specimens to be swiftly analyzed, as well as restricted time frames and limited access to taxonomic specialists.

Acknowledgements

We thank Joana Margarida Teixeira Lopes, Claude Kolwelter and Dylan Thissen for their help during the wild bee fieldwork in 2019, as well as Sophie Ogan for providing the samples from the Rhineland-Palatinate state used in this work. Nico Schneider is ac-

knowledged for his feedback on the taxonomic identification of wild bees and Gerhard Haszprunar for his collaboration. We would also like to thank Rashi Halder from the Luxembourg Centre for Systems Biomedicine in Belval for her work on the high-throughput sequencing of the library and the Aquatic Ecosystem Research Group of the University of Duisburg-Essen for their help with the bioinformatic analysis. We also thank the Zoology department at the Musée national d'histoire naturelle Luxembourg (MNHNL), especially Amanda Luttringer and Stéphanie Lippert for their useful comments on the manuscript.

Finally, we would also like to thank Christophe Praz for the constructive feedback during the revision of this manuscript. Financial support was received under the Bauer and Stemmler foundations programme "FORSCHUNGSGEIST! Next Generation Sequencing in der Oekosystemforschung". Collection permit was issued by the Ministère de l'Environnement, du Climat et du Développement durable (MECDD) Luxembourg.

References

Alberdi A, Aizpurua O, Gilbert MTP, Bohmann K (2018) Scrutinizing key steps for reliable metabarcoding of environmental samples. Methods in Ecology and Evolution 9(1): 134–147. https://doi.org/10.1111/2041-210X.12849

Amiet F, Herrmann M, Müller A, Neumeyer R (2007) Fauna Helvetica 20: Apidae 5: *Ammobates*, *Ammobatoides*, *Anthophora*, *Biastes*, *Ceratina*, *Dasypoda*, *Epeoloides*, *Epeolus*, *Eucera*, *Macropis*, *Melecta*, *Melitta*, *Nomada*, *Pasites*, *Tetralonia*, *Thyreus*, *Xylocopa*. Centre suisse de cartographie de la Faune, Neuchatel, Switzerland, 356 pp.

Amiet F, Herrmann M, Müller A, Neumeyer R (2004) Fauna Helvetica 9. Apidae 4: *Anthidium*, *Chelostoma*, *Coelioxys*, *Dioxys*, *Heriades*, *Lithurgus*, *Megachile*, *Osmia*, Stelis. Centre Suisse de Cartographie de la Faune, Neuchatel, Switzerland, 273 pp.

Amiet F, Herrmann M, Müller A, Neumeyer R (2001) Fauna Helvetica 9. Apidae 3: *Lasioglossum*, *Halictus*. Centre Suisse de Cartographie de la Faune, Neuchatel, Switzerland. 208 pp.

Amiet F, Müller A, Neumeyer R (1999) Fauna Helvetica 9. Apidae 2 : *Colletes*, *Dufourea*, *Hylaeus*, *Nomia*, *Nomioides*, *Rhophitoides*, *Rophites*, *Sphecodes*, *Systropha*. Centre Suisse de Cartographiede la Faune, Neuchatel, Switzerland, 219 pp.

Bergsten J, Bilton DT, Fujisawa T, Elliott M, Monaghan MT, Balke M, Vogler AP (2012) The effect of geographical scale of sampling on DNA barcoding. Systematic Biology 61(5): 851–869. https://doi.org/10.1093/sysbio/sys037

Biesmeijer JC, Roberts SP, Reemer M, Ohlemüller R, Edwards M, Peeters T, Settele J (2006) Parallel declines in pollinators and insect-pollinated plants in Britain and the Netherlands. Science 313(5785): 351–354. https://doi.org/10.1126/science.1127863

Brandon-Mong GJ, Gan HM, Sing KW, Lee PS, Lim PE, Wilson JJ (2015) DNA metabarcoding of insects and allies: an evaluation of primers and pipelines. Bulletin of Entomological Research 105(6): 717–727. https://doi.org/10.1017/S0007485315000681

Braukmann TW, Ivanova NV, Prosser SW, Elbrecht V, Steinke D, Ratnasingham S, Hebert PD (2019) Metabarcoding a diverse arthropod mock community. Molecular Ecology Resources 19(3): 711–727. https://doi.org/10.1111/1755-0998.13008

Brown MJ, Paxton RJ (2009) The conservation of bees: a global perspective. Apidologie 40(3): 410–416. https://doi.org/10.1051/apido/2009019

Buchner D, Leese F (2020) BOLDigger - a Python package to identify and organise sequences with the Barcode of Life Data systems. Metabarcoding and Metagenomics 4: 19–21. https://doi.org/10.3897/mbmg.4.53535

Cane JH, Sipes S (2006) Characterizing floral specialization by bees: analytical methods and a revised lexicon for oligolecty. Plant-Pollinator Interactions: From Specialization To Generalization 99: 122.

Cantú-Salazar L, Vray S, L'Hoste L, Jakubzik A, Herrera-Mesías F (2021) An addition to the list of wild bee fauna of Luxembourg: *Nomada kohli* Schmiedeknecht, 1882 (Hymenoptera, Apidae), with a list of the species of the genus *Nomada* Scopoli, 1770 recorded in the country. Bulletin de la Société des Naturalistes Luxembourgeois 123: 195–204.

Clarke LJ, Soubrier J, Weyrich LS, Cooper A (2014) Environmental metabarcodes for insects: in silico PCR reveals potential for taxonomic bias. Molecular Ecology Resources 14(6): 1160–1170. https://doi.org/10.1051/apido/2009019

Creedy TJ, Norman H, Tang CQ, Qing Chin K, Andujar C, Arribas P, Vogler AP (2020) A validated workflow for rapid taxonomic assignment and monitoring of a national fauna of bees (Apiformes) using high throughput DNA barcoding. Molecular Ecology Resources 20(1): 40–53. https://doi.org/10.1111/1755-0998.13056

Cristiano MP, Fernandes-Salomão TM, Yotoko KS (2012) Nuclear mitochondrial DNA: an Achilles' heel of molecular systematics, phylogenetics, and phylogeographic studies of stingless bees. Apidologie 43(5): 527–538. https://doi.org/10.1007/s13592-012-0122-4

Dirzo R, Young HS, Galetti M, Ceballos G, Isaac NJ, Collen B (2014) Defaunation in the Anthropocene. Science 345(6195): 401–406. https://doi.org/10.1126/science.1251817

Dogterom MH, Matteoni JA, Plowright RC (1998) Pollination of greenhouse tomatoes by the North American *Bombus vosnesenskii* (Hymenoptera: Apidae). Journal of Economic Entomology 91(1): 71–75. https://doi.org/10.1093/jee/91.1.71

Edgar RC, Flyvbjerg H (2015) Error filtering, pair assembly and error correction for next-generation sequencing reads. Bioinformatics 31(21): 3476–3482. https://doi.org/10.1093/bioinformatics/btv401

Elbrecht V, Leese F (2017a) PrimerMiner: an R package for development and in silico validation of DNA metabarcoding primers. Methods in Ecology and Evolution 8(5): 622–626. https://doi.org/10.1111/2041-210X.12687

Elbrecht V, Leese F (2017b) Validation and development of COI metabarcoding primers for freshwater macroinvertebrate bioassessment. Frontiers in Environmental Science 5: 11. https://doi.org/10.3389/fenvs.2017.00011

Elbrecht V, Leese F (2015) Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass-sequence relationships with an innovative metabarcoding protocol. PLoS ONE 10(7): 0130324. https://doi.org/10.1371/journal.pone.0130324

Elbrecht V, Hebert PDN, Steinke D (2018) Slippage of degenerate primers can cause variation in amplicon length. Scientific Reports 8: 10999. https://doi.org/10.1038/s41598-018-29364-z

Elbrecht V, Steinke D (2019) Scaling up DNA metabarcoding for freshwater macrozoobenthos monitoring. Freshwater Biology 64(2): 380–387. https://doi.org/10.7287/peerj.preprints.3456v5

Elbrecht V, Vamos EE, Steinke D, Leese F (2018) Estimating intraspecific genetic diversity from community DNA metabarcoding data. PeerJ 6: e4644. https://doi.org/10.7717/peerj.4644

Falk SJ (2015) Field guide to the bees of Great Britain and Ireland. Bloomsbury, London.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3: 294–299.

Fox J, Weisberg S, Adler D, Bates D, Baud-Bovy G, Ellison S, Graves S (2016) R. Package 'car'. Companion to applied regression. R Package version 2-1.

Goulson D, Lye GC, Darvill B (2008) Decline and conservation of bumble bees. Annual Review of Entomology 53: 191–208. https://doi.org/10.1146/annurev.ento.53.103106.093454

Gueuning M, Blaser GD, Albrecht SM, Knop E, Praz C, Frey JE (2019) Evaluating next-generation sequencing (NGS) methods for routine monitoring of wild bees: Metabarcoding, mitogenomics or NGS barcoding. Molecular Ecology Resources 19(4): 847–862. https://doi.org/10.1111/1755-0998.13013

Hebert PD, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proceedings of the National Academy of Sciences 101(41): 14812–14817. https://doi.org/10.1073/pnas.0406166101

Herrera-Mesías F, Weigand AM (2021) Updates to the checklist of the wild bee fauna of Luxembourg as inferred from revised natural history collection data and fieldwork. Biodiversity Data Journal 9: e64027. https://doi.org/10.3897/BDJ.9.e64027

Hallmann CA, Sorg M, Jongejans E, Siepel H, Hofland N, Schwan H, Goulson D (2017) More than 75 percent decline over 27 years in total flying insect biomass in protected areas. PLoS ONE, 12(10): e0185809. https://doi.org/10.1371/journal.pone.0185809

Hausmann A, Segerer AH, Greifenstein T, Knubben J, Morinière J, Bozicevic V, Habel JC (2020) Toward a standardized quantitative and qualitative insect monitoring scheme. Ecology and Evolution 10(9): 4009–4020.

Hopkins GW, Freckleton RP (2002) Declines in the numbers of amateur and professional taxonomists: implications for conservation. In Animal Conservation Forum. Cambridge University Press 5(3): 245–249. https://doi.org/10.1017/S1367943002002299

Katoh K, Misawa K, Kuma KI, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Research 30(14): 3059–3066. https://doi.org/10.1093/nar/gkf436

Klein AM, Vaissière BE, Cane JH, Steffan-Dewenter I, Cunningham SA, Kremen C, Tscharntke T (2007) Importance of pollinators in changing landscapes for world crops. Proceedings of the Royal Society B: Biological Sciences 274: 303–313. https://doi.org/10.1098/rspb.2006.3721

Leese F, Bouchez A, Abarenkov K, Altermatt F, Borja A, Bruce K, Ekrem T, Čiampor F, Čiamporová-Zaťovičová Z, Costa F, Duarte S, Elbrecht V, Fontaneto D, Franc A, Geiger M, Hering D, Kahlert M, Stroil BK, Kelly M, Keskin E, Liška I, Mergen P, Meissner K, Pawłowski J, Penev L, Reyjol Y, Rotter A, Steinke D, Wal B, Vitecek S, Zimmermann J, Weigand AM (2018) Why we need sustainable networks bridging countries, disciplines, cultures and generations for aquatic biomonitoring 2.0: a perspective derived from the

DNAqua-Net COST action. Advances in Ecological Research 58: 63–99. https://doi.org/10.1016/bs.aecr.2018.01.001

Leray M, Yang JY, Meyer CP, Mills SC, Agudelo N, Ranwez V, Boehm JT, Machida RJ (2013) A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. Frontiers in Zoology 10: 34. https://doi.org/10.1186/1742-9994-10-34

Linhart C, Shamir R (2002) The degenerate primer design problem. Bioinformatics 18: S172–S181. https://doi.org/10.1093/bioinformatics/18.suppl_1.S172

Losey JE, Vaughan M (2006) The economic value of ecological services provided by insects. Bioscience 56(4): 311–323. https://doi.org/10.1093/bioinformatics/18.suppl_1.S172

Macher TH, Beermann AJ, Leese F (2021) TaxonTableTools: A comprehensive, platform-independent graphical user interface software to explore and visualise DNA metabarcoding data. Molecular Ecology Resources 21(5): 1705–1714. https://doi.org/10.1111/1755-0998.13358

Maddison WP, Maddison DR (2019) Mesquite: a modular system for evolutionary analysis. Version 3.61 http://www.mesquiteproject.org

Magnacca KN, Brown MJ (2010) Mitochondrial heteroplasmy and DNA barcoding in Hawaiian *Hylaeus* (*Nesoprosopis*) bees (Hymenoptera: Colletidae). BMC Evolutionary Biology 10(1): 174. https://doi.org/10.1186/1471-2148-10-174

Mangiafico S, Mangiafico MS (2017) Package 'rcompanion'. Cran Repos 20: 1–71.

Marquina D, Andersson AF, Ronquist F (2019) New mitochondrial primers for metabarcoding of insects, designed and evaluated using in silico methods. Molecular Ecology Resources 19(1): 90–104.

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet Journal 17(1): 10–12. https://doi.org/10.14806/ej.17.1.200

Meusel MS, Moritz RF (1993) Transfer of paternal mitochondrial DNA during fertilization of honeybee (*Apis mellifera* L.) eggs. Current Genetics 24(6): 539–543. https://doi.org/10.1007/BF00351719

Meusnier I, Singer GA, Landry JF, Hickey DA, Hebert PD, Hajibabaei M (2008) A universal DNA mini-barcode for biodiversity analysis. BMC Genomics 9(1): 214. https://doi.org/10.1186/1471-2164-9-214

Michener CD (2007) Bees of the World. 2nd Edn. Johns Hopkins University press. Baltimore, MD.

Nieto A, Roberts SPM, Kemp J, Rasmont P, Kuhlmann M, García Criado M, Biesmeijer JC, Bogusch P, Dathe HH, De la Rúa P, De Meulemeester T, Dehon M, Dewulf A, Ortiz-Sánchez FJ, Lhomme P, Pauly A, Potts SG, Praz CQ, Window J, Michez D (2014) IUCN Global Species Programm European Red List of Bees, 1–98. https://doi.org/10.2779/77003

Piñol J, Senar MA, Symondson WO (2019) The choice of universal primers and the characteristics of the species mixture determine when DNA metabarcoding can be quantitative. Molecular Ecology 28(2): 407–419. https://doi.org/10.1111/mec.14776

Piñol J, Mir G, Gomez-Polo P, Agustí N (2015) Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. Molecular Ecology Resources 15(4): 819–830. https://doi.org/10.1111/1755-0998.12355

Piper AM, Batovska J, Cogan NO, Weiss J, Cunningham JP, Rodoni BC, Blacket MJ (2019) Prospects and challenges of implementing DNA metabarcoding for high-throughput insect surveillance. GigaScience 8(8): giz092. https://doi.org/10.1093/gigascience/giz092

Potts SG, Biesmeijer JC, Kremen C, Neumann P, Schweiger O, Kunin WE (2010a) Global pollinator declines: trends, impacts and drivers. Trends in Ecology and Evolution 25(6): 345–353. https://doi.org/10.1016/j.tree.2010.01.007

Potts SG, Roberts SP, Dean R, Marris G, Brown MA, Jones R, Settele J (2010b) Declines of managed honey bees and beekeepers in Europe. Journal of Apicultural Research 49(1): 15–22. https://doi.org/10.3896/IBRA.1.49.1.02

R Core Team (2019) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rafferty NE (2017) Effects of global change on insect pollinators: multiple drivers lead to novel communities. Current Opinion in Insect Science 23: 22–27. https://doi.org/10.1016/j.cois.2017.06.009

Rasmont P, Genoud D, Gadoum S, Aubert M, Dufrene E, Le Goff G, Gilles Mahe G, Michez D, Pauly A (2017) Hymenoptera Apoidea Gallica: liste des abeilles sauvages de Belgique, France, Luxembourg et Suisse. Edition Atlas Hymenoptera, Université de Mons, Mons, Belgium.

Rasmont P, Mersch P (1988) Première estimation de la dérive faunique chez les bourdons de la Belgique (Hymenoptera: Apidae). In Annales de la Société Royale zoologique de Belgique 118(2): 141–147.

Ratnasingham S, Hebert PD (2007) BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). Molecular Ecology Notes 7(3): 355–364. https://doi.org/10.1111/j.1471-8286.2007.01678.x

Rognes T, Flouri T, Nichols B, Quince C, Mahé F (2016) VSEARCH: a versatile open source tool for metagenomics. PeerJ 4: e2584. https://doi.org/10.7717/peerj.2584

Rubinoff D, Cameron S, Will K (2006) A genomic perspective on the shortcomings of mitochondrial DNA for "barcoding" identification. Journal of Heredity 97(6): 581–594. https://doi.org/10.1093/jhered/esl036

Scheuchl E (2006) Illustrierte Bestimmungstabellen der Wildbienen Deutschlands und Osterreichs. Apollo books, 192 pp.

Schmidt S, Schmid-Egger C, Morinière J, Haszprunar G, Hebert PD (2015) DNA barcoding largely supports 250 years of classical taxonomy: identifications for Central European bees (Hymenoptera, *Apoidea partim*). Molecular Ecology Resources 15(4): 985–1000. https://doi.org/10.1111/1755-0998.12363

Schneider N (2018) Recension douvrage: Découvrir and protéger nos abeilles sauvages. Bulletin de la Société des Naturalistes Luxembourgeois 120: 163–164.

Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012) Towards next-generation biodiversity assessment using DNA metabarcoding. Molecular Ecology 21: 2045–2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x

van der Loos LM, Nijland R (2021) Biases in bulk: DNA metabarcoding of marine communities and the methodology involved. Molecular Ecology 30(13): 3270–3288. https://doi.org/10.1111/mec.15592

Vereecken N (2018) Découvrir and protéger nos abeilles sauvages. Glénat.

Wägele H, Klussmann-Kolb A, Kuhlmann M, Haszprunar G, Lindberg D, Koch A, Wägele JW (2011) The taxonomist-an endangered race. A practical proposal for its survival. Frontiers in Zoology 8(1): 25. https://doi.org/10.1186/1742-9994-8-25

Weigand AM, Herrera-Mesías F (2020) First record of the wild bee *Eucera* (*Tetralonia*) *alticincta* (Lepeletier, 1841) in Luxembourg. Bulletin de la Société des Naturalistes Luxembourgeois 122: 141–146.

Weigand AM, Desquiotz N, Weigand H, Szucsich N (2021) Application of propylene glycol in DNA-based studies of invertebrates. Metabarcoding and Metagenomics 5: e57278. https://doi.org/10.3897/mbmg.5.57278

Weigand H, Beermann AJ, Čiampor F, Costa FO, Csabai Z, Duarte S, Ekrem T (2019) DNA barcode reference libraries for the monitoring of aquatic biota in Europe: Gap-analysis and recommendations for future work. *S*cience of the Total Environment 678: 499–524. https://doi.org/10.1016/j.scitotenv.2019.04.247

Weigand AM, Macher JN (2018) A DNA metabarcoding protocol for hyporheic freshwater meiofauna: Evaluating highly degenerate COI primers and replication strategy. Metabarcoding and Metagenomics 2: e26869. https://doi.org/10.3897/mbmg.2.26869

Westrich P, Frommer U, Mandery K, Riemann H, Ruhnke H, Saure C, Voith J (2011) Rote Liste und Gesamtartenliste der Bienen (Hymenoptera, Apidae) Deutschlands (5. Fassung, Dezember 2011) [Red List and complete species list of bees in Germany]. Bundesamt für Naturschutz (Ed.) Rote Liste der gefährdeten Tiere, Pflanzen und Pilze Deutschlands. Band 3: Wirbellose Tiere (Teil 1) [Red List of threatened animals, plants and fungi of Germany], 371–416.

Zimmermann J, Hajibabaei M, Blackburn DC, Hanken J, Cantin E, Posfai J, Evans TC (2008) DNA damage in preserved specimens and tissue samples: a molecular assessment. Frontiers in Zoology 5(1): 18. https://doi.org/10.1186/1742-9994-5-18

Zinger L, Bonin A, Alsos IG, Bálint M, Bik H, Boyer F, De Barba M (2019) DNA metabarcoding-Need for robust experimental designs to draw sound ecological conclusions. Molecular ecology 28(8): 1857–1862. https://doi.org/10.1111/mec.15060

Supplementary material 1

In silico penalty scores, barcode coverage and congruency analysis of the wild bee species of Luxembourg

Authors: Fernanda Herrera-Mesías, Imen Kharrat Ep Jarboui, Alexander M. Weigand

Data type: tables (excel file)

Explanation note: *In silico* scores, barcode coverage and species delimitation congruence of the 349 wild bee species from Luxembourg evaluated in this study.

Link: https://doi.org/10.3897/jhr.94.84617.suppl1

Supplementary material 2

Mean wild bee genus penalty scores sorted by primer pair

Authors: Fernanda Herrera-Mesías, Imen Kharrat Ep Jarboui, Alexander M. Weigand

Data type: table (excel file)

Explanation note: Mean *in silico* penalty score and transformed mean penalty score (T-Score) within each wild bee genus considered in this study. Number of MOTUs used as weights in the ANOVA analysis are also given.

Link: https://doi.org/10.3897/jhr.94.84617.suppl2

Supplementary material 3

Summary and metadata of the wild bee samples from Luxembourg and Germany used in the mock communities

Authors: Fernanda Herrera-Mesías, Imen Kharrat Ep Jarboui, Alexander M. Weigand

Data type: table (excel file)

Explanation note: Information sheet regarding the wild bee specimes used in the mock communities. Morphological identification, molecular identification, % of identity with their best BOLD match, concentration category and set-up in which they were used are included.

Link: https://doi.org/10.3897/jhr.94.84617.suppl3

Supplementary material 4

Tagged primer combinations used in the mock community experiment

Authors: Fernanda Herrera-Mesías, Imen Kharrat Ep Jarboui, Alexander M. Weigand

Data type: table (excel file)

Explanation note: Primer tags. Combinations from Elbrecht V, Leese F (2017) Validation and development of COI metabarcoding primers for freshwater macroinvertebrate bioassessment. Frontiers in Environmental Science 5: 11. Specifications about the tag combinations used in each mock community replicate are also given.

Link: https://doi.org/10.3897/jhr.94.84617.suppl4

Supplementary material 5

Overiew of MOTU data per mock community

Authors: Fernanda Herrera-Mesías, Imen Kharrat Ep Jarboui, Alexander M. Weigand

Data type: table (excel file)

Explanation note: Number of reads per MOTU found in each mock community after applying the 0.01% filters. Includes non Hymenoptera MOTUs.

Link: https://doi.org/10.3897/jhr.94.84617.suppl5

Supplementary material 6

Mock community metabarcoding results

Authors: Fernanda Herrera-Mesías, Imen Kharrat Ep Jarboui, Alexander M. Weigand

Data type: table (excel file)

Explanation note: Sequencing results of the three PCR replicates of each mock community. Only Hymenoptera MOTUs identified to species level are considered. Number of input species detected in each set-up is also given.

Link: https://doi.org/10.3897/jhr.94.84617.suppl6

Supplementary material 7

Species delimitation congruence, comparing Linnaean species assignment of the original sequences retrieved from BOLD v/s results of MOTU clustering

Authors: Fernanda Herrera-Mesías, Imen Kharrat Ep Jarboui, Alexander M. Weigand

Data type: Image (JPG file)

Explanation note: The number of species presenting incongruent clustering and the type of anomaly is shown in each case. Cases in which the incongruence may be due to sequences uploaded under incorrect species names in the database are shown in light gray.

Link: https://doi.org/10.3897/jhr.94.84617.suppl7

Supplementary material 8

***In silico* primer performance evaluation using PrimerMiner with MOTU data sorted by wild bee genus**

Authors: Fernanda Herrera-Mesías, Imen Kharrat Ep Jarboui, Alexander M. Weigand

Data type: Image (JPG file)

Explanation note: Amplification success rates are shown for each genus (dark yellow areas = successful cases with a penalty score below 100; light yellow areas = failed cases, khaki colored areas = missing information). Missing data was excluded from calculations. Mean amplification success rates based on the whole dataset are indicated at the bottom.

Link: https://doi.org/10.3897/jhr.94.84617.suppl8

Supplementary material 9

Primer pair amplification success rates based on Linnaean species

Authors: Fernanda Herrera-Mesías, Imen Kharrat Ep Jarboui, Alexander M. Weigand

Data type: Image (JPG file)

Explanation note: Squares correspond to distinct morphospecies. Dark yellow areas represent species with a penalty score below 100, light yellow areas represent species above the threshold and striped squares represent cases of discrepancy (i.e. having MOTUs in both categories). Only sequences with full length target regions were considered.

Link: https://doi.org/10.3897/jhr.94.84617.suppl9